基于大语言模型的中英文整合复杂性建模研究

李东启[1,2] 朱廷劭[1,2*]

[1]（中国科学院心理研究所, 北京 100101）

[2]（中国科学院大学, 北京 10049）

摘要 整合复杂性是心理学中用来测量个体思维结构的一个概念，主要涉及两个方面：区分性和整合性。区分性是指个体能够识别和理解信息中存在的不同观点或元素的能力；整合性是指个体能够将这些不同的观点或元素合并成一个有逻辑性和连贯性的整体的能力。整合复杂性的测量主要依靠人工对于文本内容进行分析，这些文本可以是书面材料、演讲稿、面试记录或任何其他形式的口头或书面表达。针对当前整合复杂性人工测评方法成本高、自动化评估方法精度低以及缺乏中文文本评估方案等问题，本研究基于大语言模型文本数据增强技术和模型迁移技术为整合复杂性的评估设计了对于中英文文本的自动化评估方案，并探索了整合复杂性两种子结构：精细整合复杂性和辩证整合复杂性的自动化评估方法。本文设计并实施了两个研究，首先基于大语言模型文本数据增强技术实现了对于英文文本整合复杂性的预测模型，其次基于模型迁移技术实现了对于中文文本整合复杂性的预测模型。研究结果显示：1）使用GPT-3.5-Tubo对于英文文本数据进行增强，使用预训练多语言Roberta 模型进行词向量提取，使用文本卷积神经网络模型作为下游模型。与人工标注相比，整合复杂性 Spearman 相关系数为 0.62，辩证整合复杂性相关系数为 0.51，精细整合复杂性Spearman相关系数为0.60。优于机器学习方法以及未经过数据增强的神经网络模型。2）本文在研究二中建立了与研究一中的神经网络结构一致的模型，并将研究一中最终的模型参数迁移至本研究的模型中，对于中文文本整合复杂性进行训练。在零样本的情况下，迁移学习模型整合复杂性Spearman相关系数为0.31，辩证整合复杂性Spearman相关系数为0.31，精细整合复杂性相关系数为 0.33，均优于随机参数情况下的模型表现（整合复杂性：0.17，辩证整合复杂性：0.10，精细整合复杂性：0.10）。在小样本情况下迁移学习模型整合复杂性 Spearman 相关系数为 0.73，辩证整合复杂性 Spearman 相关系数为 0.51，精细整合复杂性相关系数为 0.73。

关键词 整合复杂性，神经网络，大语言模型，迁移学习

Integrative Complexity Modeling in English and Chinese Texts based on large language model

Abstract Integrative complexity is a concept used in psychology to measure the structure of an individual's thinking in two aspects: differentiation and integration. The measurement of integrative complexity relies primarily on manual analysis of textual content, which can be written materials, speeches, interview transcript large language models, or any other form of oral or written expression. To solve the problems of high cost of manual assessment methods, low accuracy of automated assessment methods, and the lack of Chinese text assessment scheme, this study designed an automated assessment scheme for integrative complexity on Chinese and English texts. We utilized text data enhancement technique of the large language model and the model migration technique for the assessment of integrative complexity, and explored the automated assessment methods for the two sub-structures of integrative complexity, namely, the fine integration complexity and the dialectical integration complexity. In this paper, two studies are designed and implemented. Firstly, a prediction model for the integration complexity of English text is implemented based on the text data enhancement technology of large language model; secondly, a prediction model for the integration complexity of Chinese text is implemented based on the model transfer technology. The results showed that: 1) We used GPT-3.5-Tubo for English text data enhancement, a pre-trained multilingual Roberta model for word vector extraction, and a text convolutional neural network model as a downstream model. The Spearman correlation coefficient between this model's prediction of integration complexity and the manual scoring results was 0.62, with a dialectical integration complexity correlation coefficient of 0.51 and a fine integration complexity Spearman correlation coefficient of 0.60. It is superior to machine learning methods and neural network models without data enhancement. 2) In Study 2, a model with the same structure as the neural network in Study 1 was established, and the final model parameters in Study 1 were also transferred to the model in this study to train the integration complexity prediction model based on Chinese text. In the case of zero samples, the Spearman correlation coefficients of the transfer learning model for integrative complexity are 0.31, the Spearman correlation coefficient of dialectical integration complexity is 0.31, and the correlation coefficient of fine integration complexity is 0.33, all of which are better than the model in the case of random parameters (integrative complexity: 0.17, dialectical integrative complexity: 0.10, fine integrative complexity: 0.10). In the case of small samples, the Spearman correlation coefficient of the transfer learning model was 0.73, with a dialectical integration complexity correlation coefficient of 0.51 and a fine integration complexity correlation coefficient of 0.73.

Keywords Integrative Complexity, Neural Networks, Large Language Models, Transfer Learning

1 引言

整合复杂性是一种人们思考问题时的认知加工倾向，它源自个人建构理论，是一种认知风格。整合复杂性被定为可以区分不同的观点，但同时又能在一个更高的层次上，以一种连贯的方式整合这些观点的思维方式。整合复杂性的评定针对的是思维的结构或者形式，而不是思维的内容。一般认为包括区分和整合两种结构维度(Koo & Han, 2007; Levitt, 2007)。分化是指识别和理解一个问题或情境中的多个不同视角或维度的能力。它涉及对复杂情境中各种不同元素、观点或信息的辨识和区分。

有学者认为，整合复杂性并不局限于一种形式。事实上，有研究区分了承认替代观点合法性的复杂性和服务于一种主导观点的复杂性。而且，这两种类型的复杂性可以表现出明显不同的模式。研究发现，丘吉尔在第一种复杂性方面始终低于他的同事，而在第二种复杂性方面始终高于他的同事。针对出现的这种情况，Conway 等(2008)提出了一种可以对这些不同形式进行概念化和评分的整合复杂性模型。这种整合复杂性模型被称为多重复杂性模型，该模型根据 Tetlock 和 Tyler(2006)的观点，将整合复杂性分为两大类，分别是辩证整合复杂性和精细整合复杂性。辩证整合复杂性涉及对一个焦点话题相关的不同维度之间张力的隐含认识。辩证复杂性在于对复杂话题进行深入思考时，对不同维度的张力和交织进行了隐含的认识和理解，这种思考方式体现了对复杂性的接受和认识，以及在评估和决策时考虑多重因素的能力。辩证复杂性的一些标志性特征包括以下几点。首先，个体认识到一个给定观点既有好的也有不好的元素，或者多个可能相互竞争的观点都具有合理性（例如："晒太阳既有积极也有消极的方面；一方面，躺在户外晒太阳很放松，但另一方面，这也可能增加患皮肤癌的风险"）。这一点体现了对复杂问题或现象多维度、多角度的理解，即意识到在任何一个观点中都可能存在正反两面。其次，文本陈述中的限定条件也有可能反映了个体的辩证整合复杂性（如："晒太阳总体来说是不太好的，除非天气不是太热"）。这反映了在对某一观点或行为进行评估时，考虑到了条件或情境的变化，从而使得原本的观点更加具有灵活性和适应性。精细整合复杂性发生在一个单一、主导的主题以复杂的方式被论述时。精细复杂性在于对一个特定主题的深入探讨和阐述，它强调对一个观点的全面和深入的解释和支持，而不是展示和比较多个不同的观点。这种思考方式在于对一个观点的深层次挖掘和细致的阐述。详细复杂性的标志包括以下几点。首先，精细整合复杂性强调清晰区分的观点，且这些观点都与同一主题相关且具有相同的情感倾向（例如："晒太阳是个坏主意，因为它会导致皮肤癌，同时晒太阳时很难与朋友见面，从而导致孤立和抑郁"）。这一点反映了在讨论同一主题时，对单一观点的多方面、多层次的支持和发展。其次，精细整合复杂性具有一个主导的主题或观点，并为其提供了多个不同的原因或来源（如："晒太阳会导致癌症，我对此的证据来自两个非常不同的来源：科学研究和个人经验"）。这体现了对一个特定观点的深入阐述，不仅强调了观点本身，还从多个角度支持和证实这一观点。

整合复杂性作为心理学中一个重要的研究领域，拥有悠久且丰富的研究历史。它在心理学、政治学、社会学等多个学科中都有着广泛的应用和深刻的影响。整合复杂性的概念最早起源于心理学领域，最初关注的是个体在面对复杂情境时的思维方式。早期的研究主要依赖于实验室环境中的段落完成任务。在这些任务中，被试需要根据主试提供的具体情境来完成一项或几项段落的写作任务。研究人员随后根据被试的回答来评估其整合复杂性水平，主要关注被试是否能识别和整合不同的观点。在实验室环境中，研究人员通过设定特定的情境，引导被试进行思考和表达。这种方法的优点在于可以在控制的环境下观察和评估个体的思维过程。然而，这种方法也存在局限性，如可能无法完全反映个体在真实世界中的思维方式。Suedfeld 等人(1992)的研究开创性地将整合复杂性的概念应用于自然产生的档案材料，这一转变标志着整合复杂性研究方法的重大突破。他们的工作使得研究者能够分析历史人物、政治领导者、文学作品等自然产生的文本资料，从而更加深入地理解个体或群体在特定历史背景下的思维复杂性。通过分析自然产生的文本资料，研究人员可以评估个体或群体在更加广阔和真实的社会文化背景下的思维复杂性。这种方法特别适用于历史研究、政治分析和文化研究。例如，分析政治领导人的演讲和书信可以揭示他们在重大事件中的思维方式和决策过程。

尽管整合复杂性的这种评分方式在心理学研究中被广泛应用并具有重要价值，但这种评分方法在实际操作中面临着一些显著挑战。这主要有两个原因。首先，与其他更为直接和量化的分析方法不同，整合复杂性的评估涉及对文本内容的深入理解和解释。编码员需要对作者的意图和思想进行细致和微妙的推断，这要求他们不仅要理解文本的表面意义，还要深挖作者潜在的思维模式和认知结构。其次，在评估整合复杂性时，编码员常常面临着对特定语句中是否存在分化或整合的困难判断。这种判断不仅涉及对文本的字面解读，还包括对文本背后复杂思维结构的识别和评价。

鉴于评分的复杂性和主观性，编码员需要接受专门的训练才能有效地进行整合复杂性的评估。Brown 等人(1992)提供的编码手册是编码员训练的重要工具，它包含了详细的指导和案例分析，帮助编码员理解和掌握评分标准。为了确保评分的准确性和一致性，编码员在正式进行整合复杂性编码之前需要通过一系列的测试。这些测试旨在评估编码员的评分是否与编码手册中的案例达到较高的一致性。只有当编码员的评分与标准案例在一定程度上一致时，他们才被认为具备作为合格的整合复杂性编码员的资格。

虽然整合复杂性具有系统化的人工评分方法，整合复杂性评分也不可避免地面临人工评估成本过高的问题。伴随着计算机领域中机器学习方法的快速发展，整合复杂性的评估出现了更加自动化的方法。这种利用机器学习对于文本的整合复杂性进行识别的方法，可以减少人工评分过程中所需要的时间、费用和认知资源(Ambili & Rasheed, 2014)。有研究者认为，当时的机器学习模型并不能很好地对整合复杂性进行评分，自动化评估分数和人工标注方法甚至表现出没有相关性，这无疑是很差的性能(Suedfeld & Tetlock, 2014)。但是在后续的研究中自动化整合复杂性评分工具和人工标注的结果表现出了较好的相关性。

Ambili 和 Rasheed 的研究(2014)中实现了一种对于整合复杂性的较高精度的分类器，在 156 份整合复杂性文本的整合复杂性三分类任务上达到了 0.78 的准确率。Conway 等人(2014)实现的自动化整合复杂性系统在 1300 多个段落的 9 个数据集中表现出了与人类评分之间的适度的相关性，平均的相关系数达到 0.46。Dobbrick 等人(2022)实现的机器学习与深度学习回归器中，整合复杂性自动化预测任务的性能在英文社交媒体数据集上最优相关系数可以达到 0.73，在德文社交媒体数据集上的最优相关系数可以达到 0.72，与人类评分者的评分结果已经十分相近。虽然整合复杂性的自动化预测任务在当前已经十分成熟，但是目前这种自动化的评分系统只针对整合复杂性概念进行开发，且目前为止主要针对英文数据集以及德文数据集，对于多语言的整合复杂性识别以及辩证整合复杂性和精细整合复杂性自动化识别模型有所缺乏。

综上，本研究根据首先根据整合复杂性在测量领域的困难，聚焦于如何开发出高准确率的中英文整合复杂性、辩证整合复杂性以及精细整合复杂性预测模型。本研究可以分为两个子研究。研究一聚焦于开发和验证针对英文文本的整合复杂性、辩证整合复杂性和精细整合复杂性预测模型。具体而言，为了解决神经网络高数据量的要求，研究一首先探索大语言模型 GPT-3.5-Turbo 在英文文本整合复杂性上的标注性能，通过与人类标注对比评估一致性系数，判断大语言模型是否可以作为证和复杂性代理来进行整合复杂性标注任务。其次，研究一还通过大语言模型 GPT-3.5-Turbo 文本数据增强技术和文本处理技术，尝试提高模型在识别和评估英文文本中整合复杂性方面的性能。研究二通过迁移学习方法，克服语言差异带来的挑战，尝试将研究一中得到的英文文本的整合复杂性预测模型使用模型参数迁移策略应用到中文文本，通过评估迁移学习模型在零样本和少样本情况下的性能表现，为中文文本提供有效的整合复杂性评估工具，以实现对于中文文本整合复杂性、辩证整合复杂性和精细整合复杂性的高准确率评估。通过对中英文整合复杂性自动化评估工具的研发和完善，能够更好地量化和比较不同文化背景下个体的认知处理能力。这一工具不仅有助于在全球多样化的背景下评估和理解个体如何处理和整合信息，还能为心理健康、决策制定和文化适应等领域提供重要的见解。

2 研究一 基于 GPT 文本数据增强技术的英文文本整合复杂性预测模型

2.1 数据来源

本研究的分析过程使用了两个数据集，其一是整合复杂性标注手册中提供的英文文本示例以及对应的整合复杂性标注，该数据集有 155 条文本数据，但是该数据集仅包含整合复杂性标注，并没有辩证整合复杂性标注和精细整合复杂性标注；其二是本研究中经过训练的整合复杂性编码人员对于 300 条英文文本进行整合复杂性标注的数据集，该数据集包含整合复杂性、辩证整合复杂性以及精细整合复杂性的标注。其中人工标注数据集整合复杂性一致性系数为 0.75，辩证整合复杂性一致性系数为 0.73，精细整合复杂性一致性系数为 0.72。

2.2 研究方法

大语言模型在自然语言处理任务上具有优异的性能，考虑到整合复杂性标注成本高，机器学习和神经网络样本需求量大的原因，本研究首先探索大语言模型在整合复杂性标注任务上的性能，探索大语言模型能否作为人类代理来进行整合复杂性标注任务。本实验使用整合复杂性文本标注手册中提供的 155 条文本以及对应的整合复杂性标注对大预言模型的标注性能进行测试。当前大语言模型中 GPT-3.5-Turbo 被认为具有十分优良的性能，通过提示学习的方法可以完成文本分类文本回归等多种自然语言处理相关的任务，提示学习的核心思想是通过设计一个适当的提示，来引导预训练模型执行特定的任务，如分类、生成、问答等。使用相关性系数和准确率衡量整合复杂性大语言模型标注与人工标注之间的一致性，作为大语言模型整合复杂性标注性能的衡量。

整合复杂性的机器学习方法建模使用人工标注的证和复杂性数据集，该发发首先从文本中提取与证和复杂性相关的特征，这些特征可以包括词汇使用的多样句子结构的复杂度、语法模式、情感倾向等。进行特征提取之前首先要对原始文本进行分词，该数据集由于全部是英文文本，英文天然具有空格作为文本分隔符，因此可以直接使用空格对于原始文本进行分词。在特征提取阶段，我们使用 LIWC（Linguistic Inquiry and Word Count）工具提取文本中的词频特征。将提取之后的文本特征作为模型输入，使用多种模型，如线性回归、支持向量机、高斯过程回归等模型在训练集对整合复杂性进行拟合。在测试集使用相关性系数评估整合复杂性模型的性能。具体实验流程如图 2-1 所示。

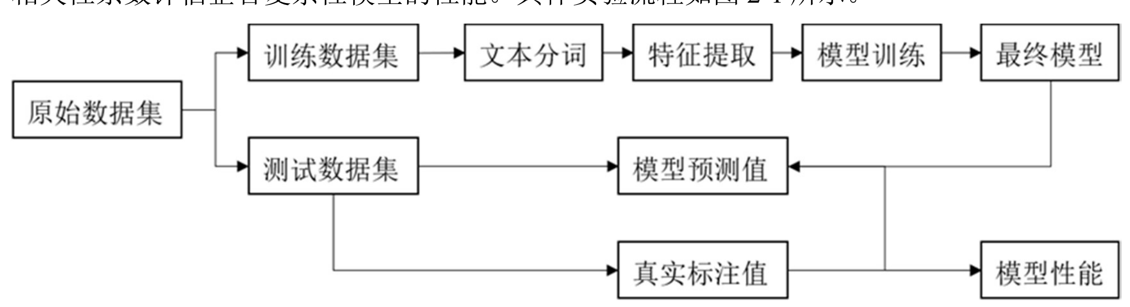


图 2-1 机器学习模型训练流程

本研究使用人工标注的证和复杂性数据集，通过预训练语言模型、词嵌入方法以及神经网络模型建立整合复杂性预测模型。本研究使用 RoBERTa（Robustly Optimized BERT Approach）多语言预训练语言模型，该模型是由 Facebook AI 研究院在 2019 年提出的一种改进版的 BERT 模型。RoBERTa 的提出主要是为了通过改进 BERT 的训练方法，进一步提高模型的性能和效率。并在下游接入 TextCNN（Text Convolutional Neural Networks）模型进行模型拟合。总体模型结构如图 2-2 所示，RoBERTa 后接入的三个 TextCNN 模型分别对整合复杂性、辩证整合复杂性和精细整合复杂性进行预测(Guo et al., 2019)。

此外，为了使得模型可以更好捕捉和优化整合复杂性、辩证整合复杂性和精细整合复杂性的概念，本研究对神经网络模型的损失函数进行针对性的设计。

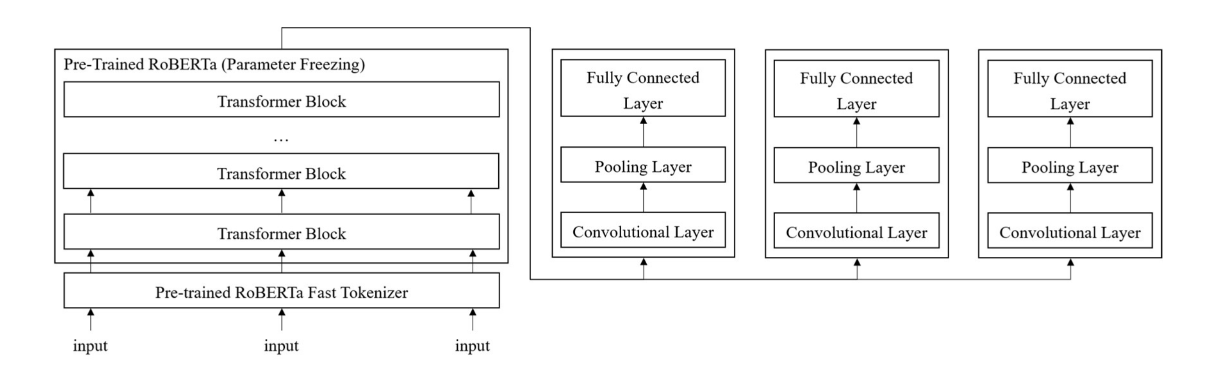


图 2-2 神经网络模型整体结构

首先，整合复杂性、辩证整合复杂性和精细整合复杂性的分值均为 1 到 7 分，其次，辩证整合复杂性得分和精细整合复杂性得分均小于等于整合复杂性得分。据此本研究设计了损失函数如图 2-3 所示，其中IC_{true}为整合复杂性人工标注分数，IC_{true}^{e}为精细整合复杂性人工标注分数，IC_{true}^{d}为辩证整合复杂性人工标注分数，IC_{pred}为整合复杂性模型预测得分，IC_{pred}^{e}为精细整合复杂性模型预测得分，IC_{pred}^{d}为辩证整合复杂性预测得分。该损失函数使用 Huber 损失作为模型主要损失函数，Huber 损失是平方损失和绝对损失的结合，可以减少对异常值的敏感性，同时保持对非异常值的高度敏感度，此外约束（2）（3）（4）使用 ReLU（Rectified Linear Unit）函数对整合复杂性、辩证整合复杂性和精细整合复杂性模型超出范围的预测值做出惩罚，如果函数的输入为正，则输出即为输入，如果输入为负，则输出零，ReLU 函数可以保证预测值尽可能位于 1 至 7 的分数范围之内。约束（4）保证辩证整合复杂性和精细整合复杂性的预测值小于等于整合复杂性的预测值。

$$\begin{aligned}
\text{Loss} = \ & \text{HuberLoss}(IC_{\text{true}} - IC_{\text{pred}}) + \text{HuberLoss}(IC^{e}_{\text{true}} - IC^{e}_{\text{pred}}) + \text{HuberLoss}(IC^{d}_{\text{true}} - IC^{d}_{\text{pred}}) && (1)\\
& + \text{ReLU}(IC_{\text{pred}} - 7) + \text{ReLU}(1 - IC_{\text{pred}}) && (2)\\
& + \text{ReLU}(IC^{e}_{\text{pred}} - 7) + \text{ReLU}(1 - IC^{e}_{\text{pred}}) && (3)\\
& + \text{ReLU}(IC^{d}_{\text{pred}} - 7) + \text{ReLU}(1 - IC^{d}_{\text{pred}}) && (4)\\
& + \text{ReLU}(IC^{e}_{\text{pred}} - IC_{\text{pred}}) + \text{ReLU}(IC^{d}_{\text{pred}} - IC_{\text{pred}}) && (5)
\end{aligned}$$

图 2-3 神经网络模型损失函数

2.2.4 基于大语言模型数据增强的整合复杂性神经网络模型建模

本研究使用为了降低整合复杂性的标注成本，使用了基于大语言模型的文本数据增强技术。传统的文本数据增强方法主要是指在深度学习普及之前，或者不直接依赖复杂模型的情况下，用于扩充或修改文本数据集的技术。比较常用的技术包括但是不限于同义词替换、随机插入、随机交换、随机删除、回译、N-gram 随机置换等方法。这些传统方法存在一定的问题，增强后的数据存在改变标签或意图的风险。基于 GPT 的文本数据增强技术利用了 GPT 模型的强大语言生成能力来自动产生新的文本数据，从而扩充和增强原始数据集。这种方法依赖于 GPT 这类大型语言模型的先进自然语言理解和自然语言生成能力，可以对原始文本进行改写，如改变句子的语态、风格或者结构，同时保持原始信息和意图不

变。与传统的文本数据增强方法相比，基于 GPT 的文本数据增强技术又更高的生成质量和多样性、更强的上下文理解能力、更强的灵活性和实用性，此外，利用 GPT 进行文本增强可以自动化地生成大量新数据，减少人工标记和校对的需要(Dai et al., 2023)。本研究要求在语义不变的情况下对原始文本的风格进行改变，来获取增强文本，具体流程如 2-4 所示。

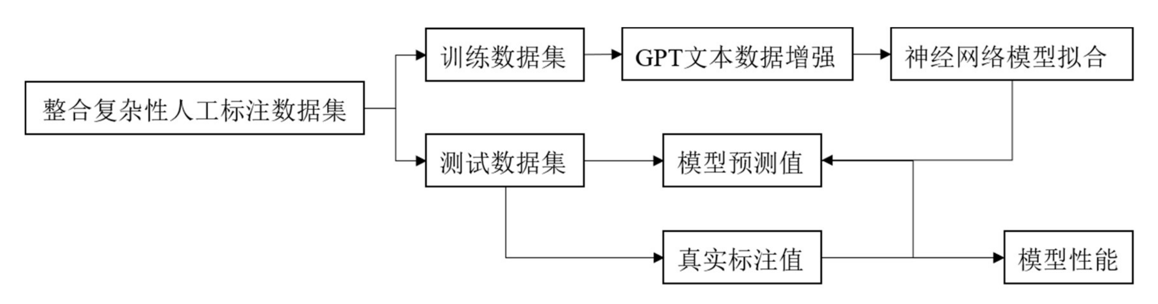


图 2-4 基于 GPT 文本数据增强技术神经网络建模

本研究中使用的神经网络的网络结构、词向量提取以及损失函数设计与未经过 GPT 数据增强技术的神经网络模型保持一致。

2.3 研究结果

本研究首先探索 GPT 模型直接进行代替人工标注直接进行整合复杂性标注的可能性。当提供给 GPT-3.5-Turbo 的提示词为整合复杂性概念和整合复杂性的分数范围时，提示词内容如图所示。使用 155 份整合复杂性标注手册提供的样例进行测试，大语言模型的整合复杂性预测值与整合复杂性人工标注之间的 Spearman 相关性系数为 0.164，Peason 相关性系数为 0.185，标注准确率为 0.219，真实值与标注值分布情况见图 2-5。

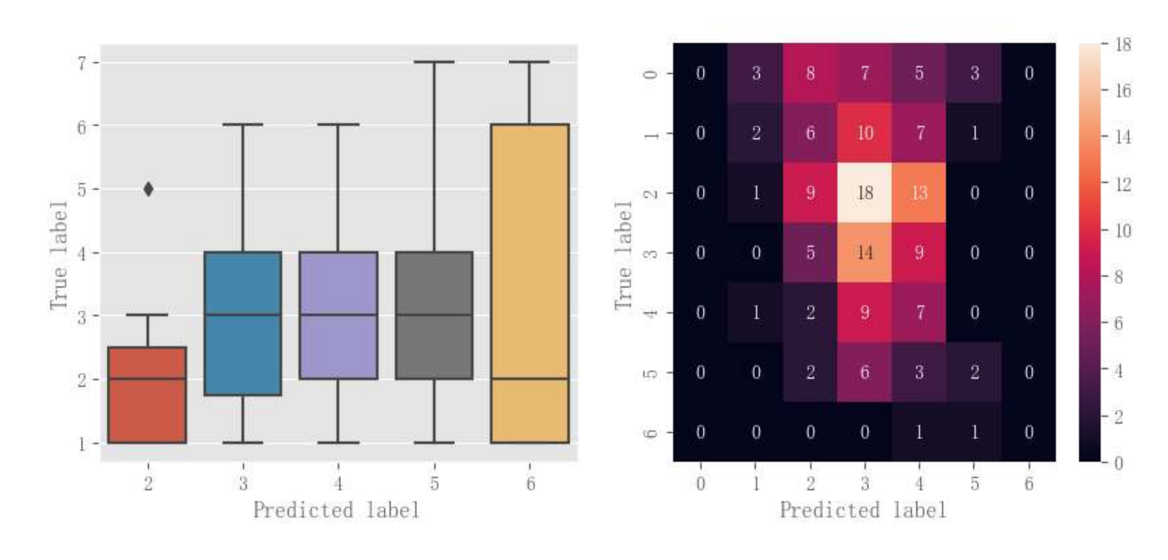


图 2-5 整合复杂性概念与分数范围提示词下预测值分布

提供给 GPT-3.5-Turbo 的提示词为整合复杂性概念、整合复杂性的分数范围和整合复杂性各个分数的含义时，提示词内容如图所示，使用 155 份整合复杂性标注手册提供的样

例进行测试。大语言模型的整合复杂性预测值与整合复杂性人工标注之间的 Spearman 相关性系数为 0.214，Peason 相关性系数为 0.244，标注准确率为 0.181，真实值与标注值分布情况见图 2-6。

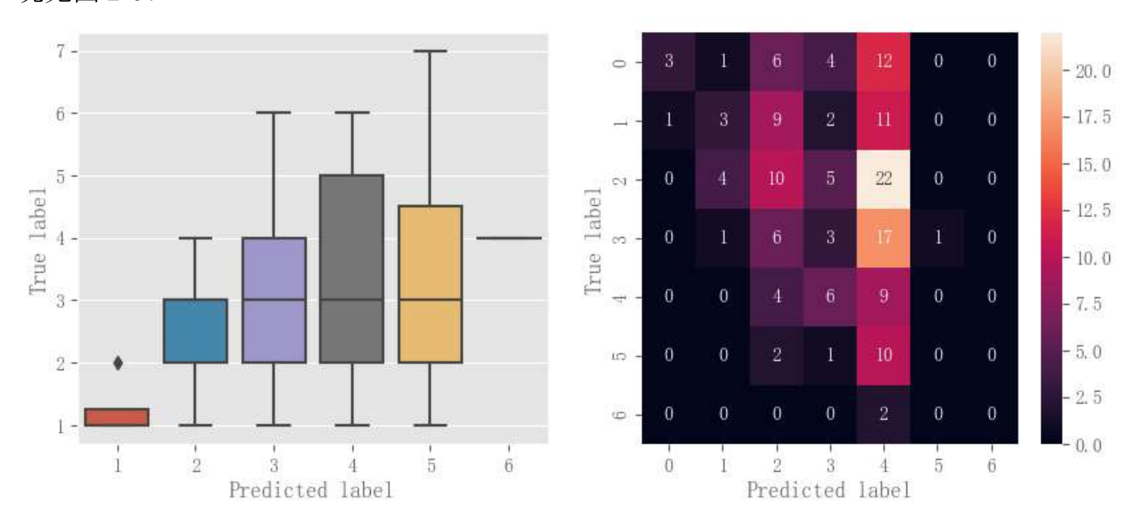


图 2-6 整合复杂性概念、分数范围和分数含义提示词下预测值分布

提供给 GPT-3.5-Turbo 的提示词为整合复杂性概念、整合复杂性的分数范围和整合复杂性各个分数的示例时，提示词内容如图所示，大使用 155 份整合复杂性标注手册提供的样例进行测试。大语言模型的整合复杂性预测值与整合复杂性人工标注之间的 Spearman 相关性系数为 0.373，Peason 相关性系数为 0.352，标注准确率为 0.297，真实值与标注值分布情况见图 2-7。

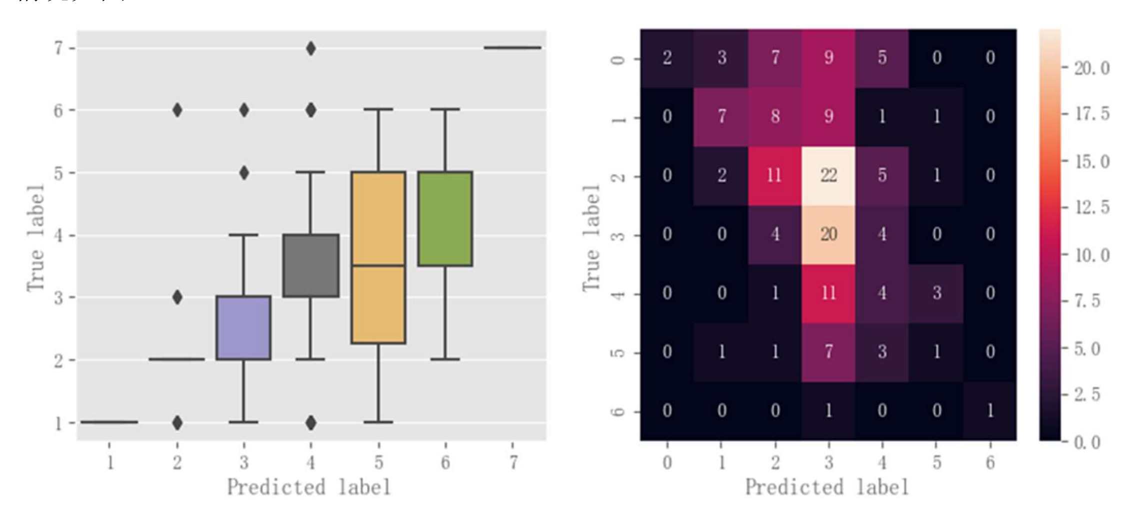


图 2-7 整合复杂性概念、分数范围和分数示例提示词下预测值分布

提供给 GPT-3.5-Turbo 的提示词为整合复杂性概念、整合复杂性的分数范围、整合复杂性各个分数的含义以及整合复杂性各个分数的示例时，使用 155 份整合复杂性标注手册提供的样例进行测试。大语言模型的整合复杂性预测值与整合复杂性人工标注之间的 Spearman 相关性系数为 0.436，Peason 相关性系数为 0.451，标注准确率为 0.219，真实值与标注值分布情况见图 2-8。

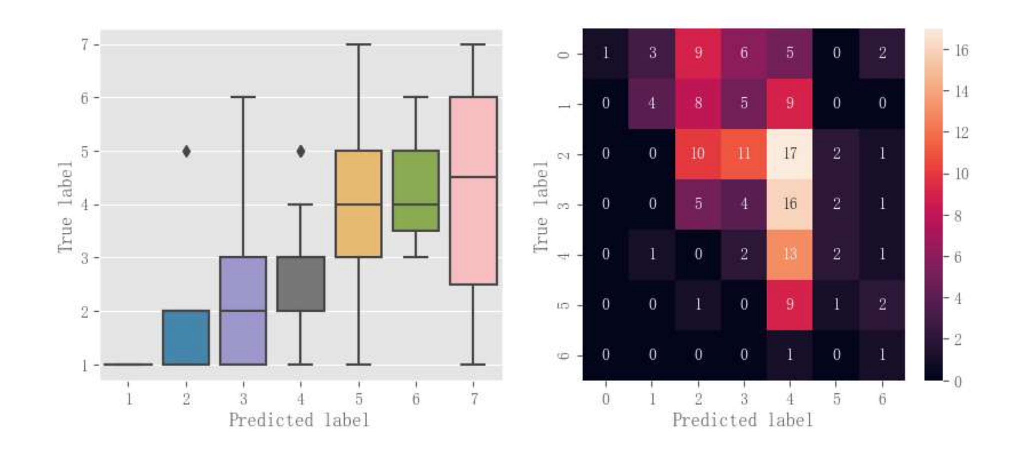


图 2-8 整合复杂性概念、分数范围、分数含义和分数示例提示词下预测值分布

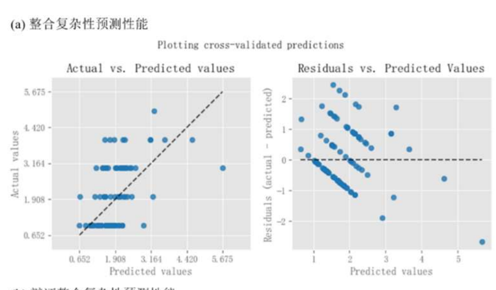
综上，当提示词的结构为整合复杂性概念、整合复杂性的分数范围、整合复杂性各个分数的含义以及整合复杂性各个分数的示例的时候，GPT-3.5-Turbo 对于整合复杂性的标注性能相较于其他结构的提示词而言是最好的，Spearman 相关性系数为 0.436，Pearson 相关性系数为 0.451，达到中等程度的相关，但是大语言模型标注的准确率较低，仅为 0.219，和人工标注结果有较大的差异，因此 GPT-3.5-Turbo 无法作为人工编码代理直接进行整合复杂性标注。

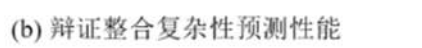
2.3.2 整合复杂性机器学习模型建模结果

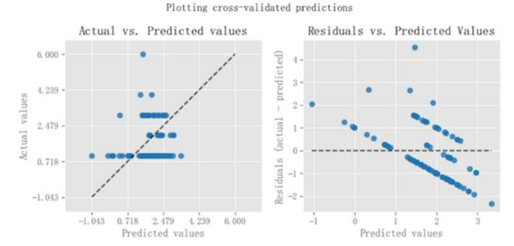
本研究使用人工标注的整合复杂性数据集，该数据集包含对于整合复杂性、辩证整合复杂性和精细整合复杂性的标注，共包含 300 份文本数据。使用 LIWC 进行模型特征提取之后，共获取 73 维特征，表 2-1 为所提取的 LIWC 特征与相对应的特征解释。

使用线性回归方法对于整合复杂性、辩证整合复杂性和精细整合复杂性进行拟合，输入特征为 LIWC 程序提取得到的 73 维心理语言学特征。线性回归是统计学中最基础且广泛使用的预测模型之一，它试图模拟`自变量和因变量之间的线性关系。使用十折交叉验证方法评估模型的预测性能，整合复杂性的 Spearman 相关性系数为 0.47，辩证整合复杂性 Spearman 相关性系数为-0.09，精细整合复杂性 Spearman 相关性系数为 0.40，真实值与预测值分布情况见图 2-9。

使用支持向量机方法对于整合复杂性、辩证整合复杂性和精细整合复杂性进行拟合，输入特征为 LIWC 程序提取得到的 73 维心理语言学特征。支持向量机是一种强大的监督学习算法，用于分类和回归分析。支持向量机的核心目标是找到一个最优的超平面（在二维空间中是一条线，在更高维度空间中是一个平面或超平面），这个超平面能够最大化地分隔不同类别的数据点。使用十折交叉验证方法评估模型的预测性能，整合复杂性的 Spearman 相关性系数为 0.48，辩证整合复杂性 Spearman 相关性系数为 0.03，精细整合复杂性 Spearman 相关性系数为 0.60，真实值与预测值分布情况见图 2-10。







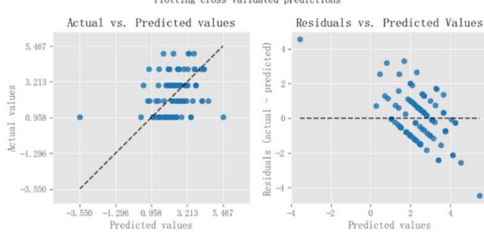
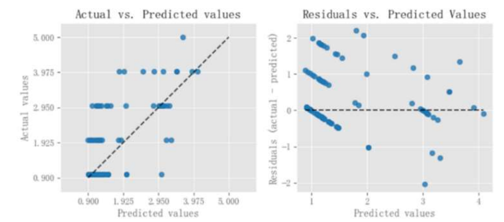
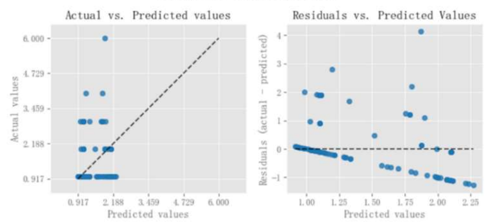


图 **2-9** 线性回归模型预测性能

(a) 整合复杂性预测性能



(b) 辩证整合复杂性预测性能



(c) 精细整合复杂性预测性能

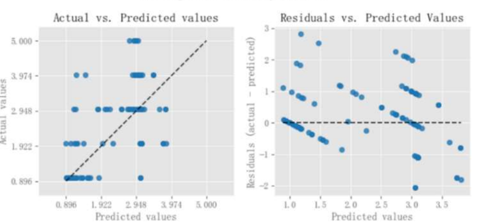


图 **2-10** 支持向量机回归模型预测性能

使用高斯过程回归方法对于整合复杂性、辩证整合复杂性和精细整合复杂性进行拟合，输入特征为 LIWC 程序提取得到的 73 维心理语言学特征。高斯过程回归，是一种贝叶斯回归技术，用于机器学习和统计学中的非线性回归问题。它通过假设数据可以由一个完全由均值函数和协方差函数定义的高斯过程生成，来进行预测和分析。高斯过程回归的核心思想是利用高斯过程来描述数据之间的关系，从而对未知点进行预测。使用十折交叉验证方法评估模型的预测性能，整合复杂性的 Spearman 相关性系数为 0.27，辩证整合复杂性 Spearman 相关性系数为-0.14，精细整合复杂性 Spearman 相关性系数为-0.11，真实值与预测值分布情况见图 2-11。

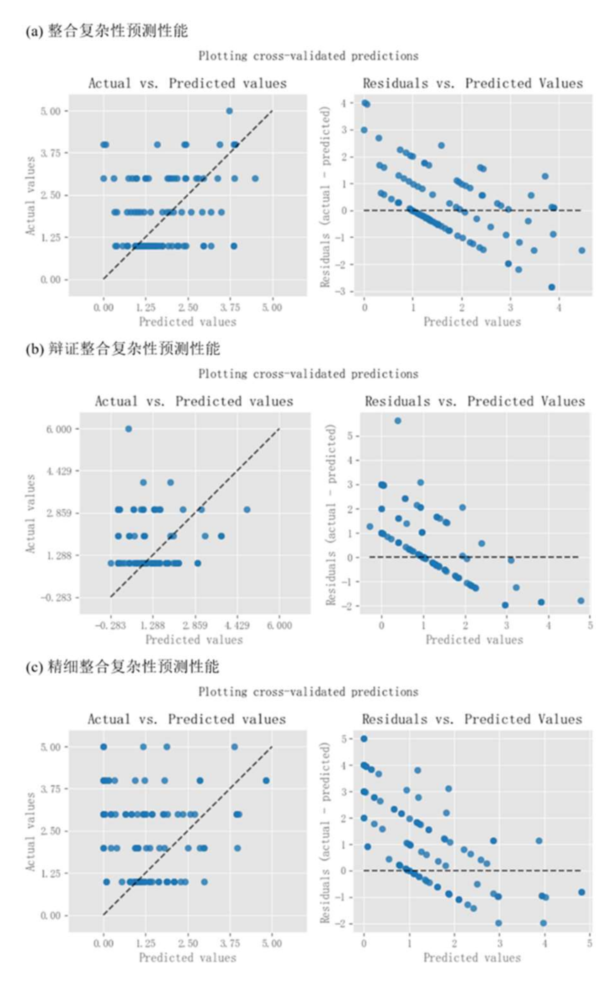
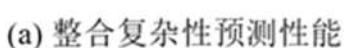


图 2-11 高斯过程回归模型预测性能

使用 XGBoost 回归方法对于整合复杂性、辩证整合复杂性和精细整合复杂性进行拟合，输入特征为 LIWC 程序提取得到的 73 维心理语言学特征。XGBoost 是一种高效且强大的梯度提升框架，旨在用于解决回归、分类和排序问题。它是基于决策树算法的集成学习方法，通过组合多个弱学习器（通常是决策树）的预测来构建一个强学习器。使用十折交叉验证方法评估模型的预测性能，整合复杂性的 Spearman 相关性系数为 0.54，辩证整合复杂性 Spearman 相关性系数为 0.14，精细整合复杂性 Spearman 相关性系数为 0.60，真实值与预测值分布情况见图 2-12。



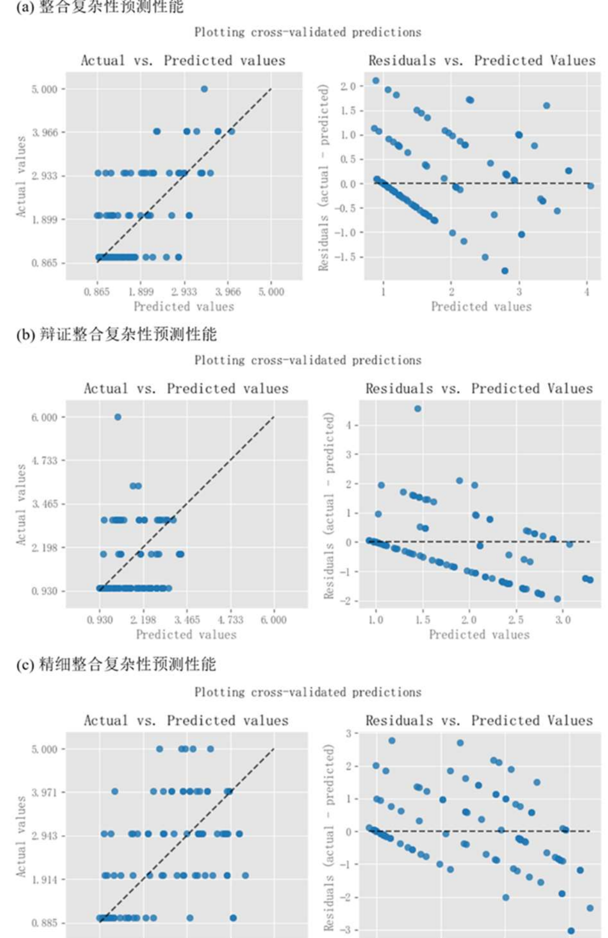


图 2-12 XGBoost 回归模型预测性能

机器学习结果汇总如表所示，对于机器学习模型而言，部分模型对于整合复杂性以及精细整合复杂性的模型性能较好，XGBoost 模型在整合复杂性上的 Spearman 相关系数可以达到 0.48，支持向量机和 XGBoost 在精细整合复杂性上的 Spearman 相关系数均可以达到 0.48，达到中等程度的相关性水平。但是目前的机器学习模型对于辩证整合复杂性的模型识别性能较差，最优模型为 XGBoost 模型，但是 Spearman 相关系数最高只有 0.16，为低相关性水平。

表 2-1 不同机器学习模型 Spearman 相关性系数

模型	整合复杂性	辩证整合复杂性	精细整合复杂性
线性回归	0.47	-0.09	0.40
支持向量机	0.48	0.03	0.60
高斯过程回归	0.27	-0.14	-0.11
XGBoost 回归	0.54	0.16	0.60

使用整合复杂性人工标注数据集进行模型拟合，数据量为 300，随机切分为训练集和测试集，其中训练集样本量为 210，测试集样本量为 90。经过 RoBERTa 预训练语言模型进

行词嵌入提取之后，每份文本被向量化为 512*768 的矩阵。使用 Adam 优化算法对于 TextCNN 网络中的模型参数进行优化，加入权重衰减项，控制模型复杂度，防止模型的过拟合情况。训练 200 个 EPOCH 之后，模型损失基本稳定。整合复杂性、辩证整合复杂性和精细整合复杂性模型预测值与人工标注得分之间的相关达到中等相关性水平。整合复杂性 Spearman 相关性系数为 0.56，辩证整合复杂性相关性系数为 0.50，精细整合复杂性相关性系数为 0.56。

使用整合复杂性人工标注数据集进行模型拟合，数据量为 300，随机切分为训练集和测试集，其中训练集样本量为 210，测试集样本量为 90。使用基于 GPT-3.5-Turbo 的文本数据增强任务之后，训练集数据量扩充为 2310 份文本，包含 2100 份增强文本和 210 份原始文本。经过 RoBERTa 预训练语言模型进行词嵌入提取之后，每份文本被向量化为 512*768 的矩阵。使用 Adam 优化算法对于 TextCNN 网络中的模型参数进行优化，加入权重衰减项，控制模型复杂度，防止模型的过拟合情况。训练 45 个 EPOCH 之后，模型损失函数基本稳定。整合复杂性、辩证整合复杂性和精细整合复杂性模型预测值与人工标注得分之间的相关达到中等相关性水平。整合复杂性 Spearman 相关性系数为 0.62，辩证整合复杂性相关性系数为 0.51，精细整合复杂性相关性系数为 0.60。

2.4 结论

本研究通过探索 GPT-3.5-Turbo 在整合复杂性标注中的性能，发现尽管其在少样本情况下表现出一定的潜力，但与人工标注的一致性有限，提示学习最优 Spearman 相关系数达到 0.44，表明大语言模型直接用于整合复杂性标注存在局限性。进一步的机器学习模型分析显示，特别是 XGBoost 回归模型在预测整合复杂性方面表现出较好的性能，整合复杂性的 Spearman 相关系数为 0.54，精细整合复杂性的 Spearman 相关系数为 0.60，辩证整合复杂性 Spearman 相关系数为 0.14。此外，本研究提出并验证了利用 GPT-3.5-Turbo 进行文本数据增强，以提高英文文本整合复杂性预测的神经网络模型性能的方法。这一方法能显著提升模型的预测准确度，英文文本神经网络模型整合复杂性 Spearman 相关性系数为 0.62，辩证整合复杂性 Spearman 相关性系数为 0.51，精细整合复杂性 Spearman 相关性系数为 0.60。

3 研究二 基于迁移学习的中文文本整合复杂性预测模型

3.1 数据来源

在本研究旨在探索基于迁移学习的中文文本整合复杂性预测模型。为了构建和验证这一模型，我们收集了来自三个不同来源的中文文本数据：微博发文文本、美团评论文本以及网络新闻文本。这些数据源被选中是因为它们覆盖了社交媒体、在线评价以及正式报道等多种文本类型，能够全面反映中文文本在不同领域中的应用和特点。该数据集包含 200 份微博发文文本，150 份美团评论文本以及 150 份网络新闻文本。其中，整合复杂性一致性

系数为 0.78，辩证整合复杂性一致性系数为 0.76，精细整合复杂性一致性系数为 0.75。

3.2 研究方法

迁移学习是机器学习领域的一个重要研究方向，它旨在将从一个任务（源任务）学习到的知识应用到另一个相关但不同的任务（目标任务）上。迁移学习的核心思想是利用已有的知识来加速和优化新任务的学习过程，特别是在新任务的标注数据较少或获取成本较高时，迁移学习能够显著提高学习效率和性能。本研究使用了和研究一相同的方法，采用了相同的词嵌入技术和模型架构来对中文文本的整合复杂性进行深入建模。在这一过程中，我们特别关注了模型参数迁移这一策略的应用效果。具体而言，我们将之前针对英文文本整合复杂性预测任务中训练好的模型参数，直接迁移到新的中文文本整合复杂性预测模型中。这一步骤旨在探索通过迁移学习在不同语言间共享知识的可行性及其对模型性能的影响。为了深入评估模型参数迁移的效果，我们对比了两种不同初始化方式下模型的性能：一种是使用随机初始化的模型参数，另一种则是采用从英文模型中迁移得来的参数进行初始化，通过评估这两种模型在样本量为 150 的测试集上的性能来判断零样本迁移学习模型对于中文文本整合复杂性预测的优势。

在传统的迁移学习模型基础上，少样本迁移学习（Few-Shot Learning）模型进一步针对在目标任务上只有极少量标注数据可用的情况进行优化。少样本迁移学习有元学习、模型微调以及数据增强等多种学习策略。其中，元学习通过在多个任务上训练，学习到一种学习策略，使模型能够在遇到新任务时快速适应。模型微调是指在源任务上预训练一个模型，然后使用少量目标任务数据对模型进行微调，以达到良好的泛化能力(Zhuang et al., 2020)。数据增强是指通过生成式模型或其他技术手段增加目标任务的训练样本，降低过拟合风险。

本研究尝试将少样本迁移学习模型应用于中文文本整合复杂性的预测任务中，探索在极少标注数据情况下实现准确预测的可能性。我们首先在丰富的英文文本整合复杂性标注数据集上训练了基模型，以获取广泛的语言理解和文本分析能力。随后，总中文标注数据集中选取了 350 份中文文本整合复杂性数据作为目标任务的训练集，使用 150 份中文文本整合复杂性数据作为目标任务的测试集，评估少样本迁移学习模型与未迁移学习模型在中文文本整合复杂性预测任务上的性能差异。

3.3 研究结果

根据表 3-1 所展示的实验结果数据，可以看到零样本迁移学习模型在整合复杂性、辩证整合复杂性和精细整合复杂性的评估指标上均显著优于非迁移学习模型。具体而言，对于整合复杂性的预测，非迁移学习模型的 Spearman 相关性系数为 0.17，而迁移学习模型的 Spearman 相关性系数提高到了 0.31。在辩证整合复杂性的预测中，迁移学习模型的 Spearman 相关性系数从非迁移学习模型的 0.06 提升至 0.31，展现出了更加显著的提升。同

样，在精细整合复杂性的评估上，迁移学习模型的 Spearman 相关性系数达到了 0.33，相比非迁移学习模型的 0.10 有了显著的提高。

这些结果表明，相较于非迁移学习模型，采用迁移学习方法能够显著提高模型在预测文本整合复杂性方面的性能。迁移学习模型能够有效地利用从源任务中学习到的知识，即使在目标任务上没有经过特别训练，也能达到较高的预测准确度。尽管源任务（英文文本整合复杂性预测）与目标任务（中文文本整合复杂性预测）在语言特性上存在显著差异，迁移学习模型依然能够利用在源任务中学习到的知识，有效地提升在目标任务上的预测能力。这表明了语言之间在表示文本复杂性方面可能存在的共性，以及迁移学习在捕捉这些共性方面的潜力。

表 3-2 零样本迁移学习模型与非迁移学习模型效果评估

	整合复杂性	辩证整合复杂性	精细整合复杂性
非迁移学习模型	0.17	0.06	0.10
迁移学习模型	0.31	0.31	0.33

采用少样本迁移学习的模型在整合复杂性、辩证整合复杂性以及精细整合复杂性的预测任务上均展现出了相较于非迁移学习模型更优的性能。具体而言，非迁移学习模型在整合复杂性的 Spearman 相关性系数为 0.68，迁移学习模型在整合复杂性的 Spearman 相关性系数为 0.73；非迁移学习模型在辩证整合复杂性的 Spearman 相关性系数为 0.47，迁移学习模型在辩证整合复杂性的 Spearman 相关性系数为 0.51；非迁移学习模型在精细整合复杂性的 Spearman 相关性系数为 0.72，迁移学习模型在精细整合复杂性的 Spearman 相关性系数为 0.73。

表 3-3 少样本迁移学习模型与非迁移学习模型效果评估

	整合复杂性	辩证整合复杂性	精细整合复杂性
非迁移学习模型	0.68	0.47	0.72
迁移学习模型	0.73	0.51	0.73

3.4 结论

根据研究结果，零样本和少样本迁移学习模型在预测整合复杂性、辩证整合复杂性和精细整合复杂性方面显著优于传统非迁移学习模型。这强调了迁移学习在处理跨语言文本分析任务中的有效性和潜力，尤其是在面对数据稀缺的情况下。少样本迁移学习模型的结果进一步证实了在有限样本情况下，迁移学习能够提升模型的预测准确度，展示了迁移学习策略在提高模型泛化能力方面的优势。中文文本的迁移学习模型在零样本情况下，整合复杂性 Spearman 相关系数为 0.31，精细整合复杂性 Spearman 相关系数为 0.33，辩证整合复杂性 Spearman 相关系数为 0.31；在少样本情况下，迁移学习模型整合复杂性 Spearman

相关系数为 0.73，精细整合复杂性 Spearman 相关系数为 0.73，辩证整合复杂性 Spearman 相关系数为 0.51。这些发现为跨语言和跨领域的文本复杂性研究提供了重要见解，并为未来的迁移学习应用和研究开辟了新路径。此外，本研究开发出了高精度的适用于中文文本的整合复杂性预测模型，实现了对于整合复杂性、辩证整合复杂性和精细整合复杂性的高准确率预测。

4 总讨论

本项研究旨在探讨不同类型整合复杂性在中英文文本建模的性能效。通过采用大型语言模型 GPT-3.5-Turbo 进行文本数据增强技术，结合社交网络平台的大数据资源，本研究展开了系统性探索。研究的第一部分着重于评估 GPT-3.5-Turbo 在标注整合复杂性方面的性能表现。继而，本研究进一步探讨了机器学习模型和神经网络模型在预测英文文本中整合复杂性、精细整合复杂性及辩证整合复杂性方面的能力。此外，本研究提出了一种借助 GPT-3.5-Turbo 进行文本数据增强的新方法，并使用增强后的数据对英文文本中的整合复杂性进行再次建模，成功构建了一种具有较高识别性能的神经网络模型。在研究的第二部分，通过应用迁移学习技术，开发了一种适用于中文文本的整合复杂性预测模型，实现了在中文语境下对整合复杂性、辩证整合复杂性和精细整合复杂性的高准确率预测。本研究为整合复杂性的测量提供了新的视角和技术支持。

4.1 大语言模型进行数据标注的潜力

研究一深入探讨了利用 GPT-3.5-Turbo 等先进大型语言模型在少样本学习情境下处理复杂自然语言处理标注任务，特别是整合复杂性标注任务的能力。研究结果显示，尽管大型语言模型在广泛的 NLP 应用中展现了显著优良的性能，但在与人类专家相比较时，其在整合复杂性标注任务上的性能仍显不足。我们认为这可能有两个主要的原因。首先，GPT-3.5-Turbo 及类似模型具有性能上的局限性，在性能上的局限部分源自其模型的设计和模型的训练参数规模。尽管这些模型具有庞大的参数规模和广泛的知识库，它们在处理长篇复杂文本信息，尤其是需要深层次理解和整合的场景中的能力，虽然和其他模型相比具有无可比拟的优势，但是对于某些复杂概念理解有待加强(Brown et al., 2020)。模型当前的处理能力受到其训练机制和算法优化的限制，导致在复杂标注任务中的表现不尽人意。其次，与简单的情感极性标注任务相比，整合复杂性标注对模型的要求可能超出了现有能力范围，整合复杂性的标注甚至对于专业的标注人员而言都是一种比较困难的任务。这种任务往往需要对长篇文本进行深入理解、跨文本信息整合及高度推理，这些要求对大型语言模型在语义理解、上下文关联及信息整合方面的能力提出了挑战(Radford et al., 2019)。

尽管存在上述挑战，大型语言模型的快速进展预示着未来在处理复杂自然语言处理任务上的性能可能会显著提升。为克服当前限制，未来模型可能通过采用更先进的训练策略、

复杂的模型架构和精细化优化算法来实现。例如，通过引入更精细的注意力机制、改进的知识蒸馏技术和更有效的数据预处理方法，可以增强模型在处理复杂文本及执行高级认知任务上的能力。此外，未来研究可以探索大语言模型在一些简单任务上的标注性能，比如情感极性标注等，以探索大语言模型在类似任务上的潜力(Han et al., 2021)。

4.2 神经网络在心理变量自动化识别中的优劣势

此外，本研究比较了传统机器学习方法和基于神经网络的方法在整合复杂性预测任务上的性能，发现虽然两类方法在预测整合复杂性和精细整合复杂性方面性能相当，但神经网络模型在处理辩证整合复杂性任务上表现出明显优势。这一差异突显了两种方法在信息处理和知识表达上的本质不同。神经网络通过学习文本中的词嵌入和深层结构，能够抽象出更复杂的知识表示，为理解文本背后的逻辑和上下文信息提供了强大的能力。这使得神经网络在整合复杂性的多个维度，尤其是在处理辩证整合复杂性方面，展现出更为卓越的性能(LeCun et al., 2015)。传统机器学习模型的性能表现确实在很大程度上依赖于特征提取。在传统机器学习流程中，特征提取是一个关键步骤，它涉及将原始数据转化为模型能够理解和处理的格式。这一步骤通常需要领域知识和专业技能，以确定哪些特征对于预测任务最为重要(Jordan & Mitchell, 2015)。LIWC 特征可能无法提取和学习到辩证整合复杂性的所需要的特征，这导致辩证整合复杂性在任何传统机器学习模型上均表现出比较差的性能。此外，传统机器学习方法聚焦于对于单个心理变量的自动化识别，这种方法不能有效利用多个心理学变量之间的关系，但是神经网络方法不仅仅可以对单个心理学变量进行更高准确率的建模，还可以同时利用关联的心理学变量的信息，对多个心理学变量同时进行建模，并通过约束关系尽可能使模型预测值保持和真实值相同的关联关系(Radford et al., 2019)。

综上所述，本研究不仅揭示了大型语言模型在复杂自然语言处理标注任务中的应用潜力和局限性，也为未来研究提供了改进方向和策略，以实现更精准和高效的整合复杂性建模和预测。此外，通过探索和优化大型语言模型与专家系统的结合应用，有望进一步扩展模型的应用范围，使其更好地适应复杂和专业化的自然语言处理标注任务。

4.3 迁移学习模型在心理学变量自动化识别中的应用

研究二深入探讨了迁移学习技术在中文文本整合复杂性、辩证整合复杂性以及精细整合复杂性预测模型构建中的应用及其效果。通过采用模型参数迁移策略，该研究成功地建立了能够高准确度预测上述复杂性指标的模型，并特别强调了迁移学习在零样本及少样本学习场景下的性能表现。这一发现不仅彰显了迁移学习技术在促进模型适应性和泛化能力方面的潜力，也揭示了迁移学习模型能够捕捉到跨文化背景下整合复杂性概念的通用知识表达结构(Marcelino, 2018)。迁移学习模型在建模过程中成功捕获了反映不同文化背景下整合复杂性的通用知识结构。这表明，尽管整合复杂性的具体表达可能受到语言和文化的影响，但存在一定的概念和结构是普遍适用的(Suedfeld & Tetlock, 2014)。这一发现为理解整

合复杂性在全球范围内的应用提供了有价值的视角，同时也为跨文化研究提供了新的方法学工具。在零样本和少样本情况下，迁移学习模型展现出显著的性能优势，证明了其在处理数据稀缺情况下的强大能力(Beltagy et al., 2022)。然而，随着样本量的增加，本研究指出迁移学习模型可能不会稳定保持性能优势(Kaplan et al., 2020)。这一观察暗示，在充足数据支持的情况下，从头训练的模型可能通过学习任务特定的细节来达到或超越迁移学习模型的性能。这一点突出了在实际应用中选择模型策略时需要考虑的权衡和挑战，尤其是在评估数据可用性和模型训练成本时。

未来的研究应进一步探索迁移学习在不同自然语言处理任务中的应用潜力，尤其是在跨文化和跨语言环境下的复杂性建模。此外，研究应当探讨改进迁移学习模型的策略，以保持其在样本量大时的性能优势，例如通过开发更精细的微调方法或结合迁移学习与最新的模型训练技术。同时，研究还应该关注如何量化和优化迁移学习模型在不同文化背景下的适用性和效率，以及如何利用迁移学习捕获和利用跨文化的共性知识，为全球范围内的自然语言处理应用提供支持。

5 总结

本研究聚焦于实现对于中英文文本高准确率整合复杂性自动化识别模型的建立，并在此基础上探索了在社交网络媒体表达中整合复杂性对于自杀意念的影响作用，以及自杀行为发生的微博用户在自杀行为发生前夕网络媒体表达中的整合复杂性变化趋势。在此过程中，我们利用大语言模型文本数据增强技术和迁移学习方法克服了传统神经网络方法数据量要求大的困难，此外，还探索了大语言模型作为编码代码在整合复杂性标注上的标注性能。研究结论如下：

（1）首先探索了大语言作为整合复杂性标注代理进行整合复杂性标注的可能性，大语言模型在少样本提示学习的情况下，最优的模型性能可以达到 0.45，表现出了一定的标注潜力，但是无法达到人类标注水平；使用基于大语言模型的文本数据增强技术建立了针对英文文本的整合复杂性自动化识别工具，该自动化识别模型可以识别英文文本的整合复杂性、辨证整合复杂性和精细整合复杂性。经过测试集性能验证，该自动化识别模型在整合复杂性上与人工标注的 Spearman 相关性达到 0.62，在辩证整合复杂性上与人工标注的 Spearman 相关性达到 0.51，在精细整合复杂性上与人工标注的 Spearman 相关性达到 0.60。

（2）使用模型参数迁移的方法，将研究一中建立的英文整合复杂性模型进行模型参数迁移，从而建立了中文整合复杂性的高准确率识别模型，该自动化识别模型可以识别中文文本的整合复杂性、辩证整合复杂性和精细整合复杂性。经过测试集性能验证，该自动化识别模型在整合复杂性上与人工标注的 Spearman 相关性达到 0.73，在辩证整合复杂性上与人工标注的 Spearman 相关性达到 0.51，在精细整合复杂性上与人工标注的 Spearman 相关性达到 0.73，并且证明在零样本和小样本情况下，迁移学习模型均表现一定的性能优势。

参考文献

Ambili, A. K., & Rasheed, K. M. (2014, December). Automated scoring of the level of integrative complexity from text using machine learning. In 2014 13th International Conference on Machine Learning and Applications (pp. 300-305). IEEE.

Baker-Brown, G., Ballard, E. J., Bluck, S., De Vries, B., Suedfeld, P., & Tetlock, P. E. (1992). The conceptual/integrative complexity scoring manual. Motivation and personality: Handbook of thematic content analysis, 401-418.

Beltagy, I., Cohan, A., Logan IV, R., Min, S., & Singh, S. (2022, May). Zero-and few-shot nlp with pretrained language models. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts (pp. 32-37).

Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. In Proceedings of the 58th annual meeting of the association for computational linguistics (pp. 5185-5198).

Brewer, M. B., & Chen, Y. R. (2007). Where (who) are collectives in collectivism? Toward conceptual clarification of individualism and collectivism. Psychological review, 114(1), 133.

Brodbeck, F. C., Kugler, K. G., Fischer, J. A., Heinze, J., & Fischer, D. (2021). Group-level integrative complexity: Enhancing differentiation and integration in group decision-making. GROUP PROCESSES & INTERGROUP RELATIONS, 24(1), 125-144.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. Advances in neural information processing systems, 33, 1877-1901.

Brundidge, J., Reid, S. A., Choi, S., & Muddiman, A. (2014). The "Deliberative Digital Divide:" Opinion Leadership and Integrative Complexity in the US Political Blogosphere. POLITICAL PSYCHOLOGY, 35(6), 741-755.

Burge, T. (1986). Individualism and psychology. The Philosophical Review, 95(1), 3-45.

Carroll, J., & Bright, A. D. (2009). The Integrative Complexity of Wildfire Management Scale: Are We There Yet? HUMAN ECOLOGY REVIEW, 16(2), 211-221.

Conway, L. G., Harris, K. J., Catley, D., Gornick, L. J., Conway, K. R., Repke, M. A., & Houck, S. C. (2017). Cognitive complexity of clients and counsellors during motivation-based treatment for smoking cessation: An observational study on occasional smokers in a US college sample. BMJ Open, 7(10).

Conway, L. G., III, Conway, K. R., Gornick, L. J., & Houck, S. C. (2014). Automated Integrative Complexity. POLITICAL PSYCHOLOGY, 35(5), 603-624.

Conway, L. G., III, Gornick, L. J., Burfeind, C., Mandella, P., Kuenzli, A., Houck, S. C., & Fullerton, D. T. (2012). Does Complex or Simple Rhetoric Win Elections? An Integrative Complexity Analysis of U.S. Presidential Campaigns. POLITICAL PSYCHOLOGY, 33(5), 599-618.

Conway, L. G., III, & Zubrod, A. (2022). Are US Presidents Becoming Less Rhetorically Complex? Evaluating the Integrative Complexity of Joe Biden and Donald Trump in Historical Context. JOURNAL OF LANGUAGE AND SOCIAL PSYCHOLOGY, 41(5), 613-625.

Conway, L. G., Thoemmes, F., Allison, A. M., Towgood, K. H., Wagner, M. J., Davey, K., Salcido, A., Stovall, A. N., Dodds, D. P., Bongard, K., & Conway, K. R. (2008). Two ways to be complex and why they matter: Implications for attitude strength and lying. Journal of Personality and Social Psychology, 95(5), 1029–1044.

Dai, H., Liu, Z., Liao, W., Huang, X., Cao, Y., Wu, Z., Zhao, L., Xu, S., Liu, W., Liu, N., Li, S., Zhu, D., Cai, H., Sun, L., Li, Q., Shen, D., Liu, T., & Li, X. (2023). AugGPT: Leveraging ChatGPT for Text Data Augmentation. arXiv preprint arXiv:2302.13007, 10.

Davidson, K., Livingstone, S., McArthur, K., Dickson, L., & Gumley, A. (2007). An integrative complexity analysis of cognitive behaviour therapy sessions for borderline personality disorder. PSYCHOLOGY AND PSYCHOTHERAPY-THEORY RESEARCH AND PRACTICE, 80, 513–523.

Dobbrick, T., Jakob, J., Chan, C.-H., & Wessler, H. (2022). Enhancing Theory-Informed Dictionary Approaches with “Glass-box” Machine Learning: The Case of Integrative Complexity in Social Media Comments. COMMUNICATION METHODS AND MEASURES, 16(4), 303-320.

Freud, S. & others. (1917). Mourning and melancholia. The standard edition of the complete psychological works of Sigmund Freud, 14(1914–1916), 237–258.

Gregory, A. L., & Piff, P. K. (2021). Finding uncommon ground: Extremist online forum engagement predicts integrative complexity. PLOS ONE, 16(1), e0245651.

Guo, B., Zhang, C., Liu, J., & Ma, X. (2019). Improving text classification with weighted word embeddings via a multi-channel TextCNN model. Neurocomputing, 363, 366–374.

Hagtvedt, H., & Vohs, K. D. (2017). Art enhances meaning by stimulating integrative complexity and aesthetic interest. BEHAVIORAL AND BRAIN SCIENCES, 40.

Han, X., Zhang, Z., Ding, N., Gu, Y., Liu, X., Huo, Y., ... & Zhu, J. (2021). Pre-trained models: Past, present and future. AI Open, 2, 225-250.

Hassan, O., & Featherstone, C. (2021). Trump’s low conceptual complexity leadership and the vanishing “unpredictability doctrine”. CAMBRIDGE REVIEW OF INTERNATIONAL AFFAIRS, 34(3), 348-359.

Hedderich, M. A., Lange, L., Adel, H., Strötgen, J., & Klakow, D. (2020). A survey on recent approaches for natural language processing in low-resource scenarios. arXiv preprint arXiv:2010.12309.

Hoffman, L. H., & Slater, M. D. (2007). Evaluating public discourse in newspaper opinion articles:

Values-framing and integrative complexity in substance and health policy issues. JOURNALISM & MASS COMMUNICATION QUARTERLY, 84(1), 58-74.

Houck, S. C., & Conway, L. G., III. (2019). Strategic Communication and the Integrative Complexity-Ideology Relationship: Meta-Analytic Findings Reveal Differences Between Public Politicians and Private Citizens in Their Use of Simple Rhetoric. POLITICAL PSYCHOLOGY, 40(5), 1119-1141.

Huang, F., Li, S., Li, D., Yang, M., Ding, H., Di, Y., & Zhu, T. (2022). The Impact of Mortality Salience, Negative Emotions and Cultural Values on Suicidal Ideation in COVID-19: A Conditional Process Model. International Journal of Environmental Research and Public Health, 19(15), 9200.

Huang, M., Zhu, X., & Gao, J. (2020). Challenges in building intelligent open-domain dialog systems. ACM Transactions on Information Systems (TOIS), 38(3), 1-32.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

Kahn, J. H., Tobin, R. M., Massey, A. E., & Anderson, J. A. (2007). Measuring emotional expression with the Linguistic Inquiry and Word Count. The American journal of psychology, 120(2), 263-286.

Kalyan, K. S., Rajasekharan, A., & Sangeetha, S. (2021). Ammus: A survey of transformer-based pretrained models in natural language processing. arXiv preprint arXiv:2108.05542.

Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., ... & Amodei, D. (2020). Scaling laws for neural language models. arXiv preprint arXiv:2001.08361.

Kenton, J. D. M.-W. C., & Toutanova, L. K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. Proceedings of naacL-HLT, 1, 2.

Koo, J., & Han, J. (2007). Integrative complexity in correspondence. JOURNAL OF PACIFIC RIM PSYCHOLOGY, 1(2), 60-69.

Koo, J., Han, J., & Kim, J. (2002). Integrative complexity of South-North Korean correspondences—A Time-series analysis, 1984-1994. JOURNAL OF CONFLICT RESOLUTION, 46(2), 286-304.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

Levitt, J. M. (2007). Integrative complexity: An important theory for peace psychologists. JOURNAL OF PACIFIC RIM PSYCHOLOGY, 1(2), 70-72.

Liu, P., Wang, X., Xiang, C., & Meng, W. (2020). A survey of text data augmentation. 2020 International Conference on Computer Communication and Network Security (CCNS), 191–195.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., &

Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.

Marcelino, P. (2018). Transfer learning from pre-trained models. Towards data science, 10, 23.

Matsumoto, D. (1999). Culture and self: An empirical assessment of Markus and Kitayama's theory of independent and interdependent self-construals. Asian Journal of Social Psychology, 2(3), 289-310.

Matsumoto, D., Yoo, S. H., & Nakagawa, S. (2008). Culture, emotion regulation, and adjustment. Journal of personality and social psychology, 94(6), 925.

McHaney, R., Tako, A., & Robinson, S. (2018). Using LIWC to choose simulation approaches: A feasibility study. Decision Support Systems, 111, 1-12.

Min, B., Ross, H., Sulem, E., Veyseh, A. P. B., Nguyen, T. H., Sainz, O., ... & Roth, D. (2023). Recent advances in natural language processing via large pre-trained language models: A survey. ACM Computing Surveys, 56(2), 1-40.

Mylek, M. R., & Schirmer, J. (2019). Thinking about Fuel Management: The Potential of Integrative Complexity Theory to Inform Design of Communication about Fuel Management Used To Reduce Wildfire Risk. SOCIETY & NATURAL RESOURCES, 32(9), 983-1002.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. Psychological review, 108(2), 291.

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. Psychological bulletin, 128(1), 3.

Park, G., & DeShon, R. P. (2018). Effects of group-discussion integrative complexity on intergroup relations in a social dilemma. ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES, 146, 62–75.

Qiu, X., Sun, T., Xu, Y., Shao, Y., Dai, N., & Huang, X. (2020). Pre-trained models for natural language processing: A survey. Science China Technological Sciences, 63(10), 1872-1897.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI blog, 1(8), 9.

Savage, S., Liht, J., & Williams, R. (2009). Being Muslim Being British: Preventing Extremist Violence Through Raising Integrative Complexity (pp. 80-94). IOS Press.

Shinde, P. P., & Shah, S. (2018). A review of machine learning and deep learning applications. 2018 Fourth international conference on computing communication control and automation (ICCUBEA), 1–6.

Stickland, A. C., & Murray, I. (2019). Bert and pals: Projected attention layers for efficient adaptation in multi-task learning. In International Conference on Machine Learning (pp. 5986-

5995). PMLR.

Suedfeld, P., DeVries, B., Bluck, S., Wallbaum, A., & Schmidt, P. (1996). Intuitive perceptions of decision-making strategy: Naive assessors' concepts of integrative complexity. INTERNATIONAL JOURNAL OF PSYCHOLOGY, 31(5), 177-190.

Suedfeld, P., & Leighton, D. (2002). Early communications in the war against terrorism: An integrative complexity analysis. POLITICAL PSYCHOLOGY, 23(3), 585-599.

Suedfeld, P., & Tetlock, P. E. (2014). Integrative Complexity at Forty: Steps Toward Resolving the Scoring Dilemma. POLITICAL PSYCHOLOGY, 35(5), 597-601.

Suedfeld, P., Tetlock, P. E., & Ramirez, C. (1977). War, Peace, and Integrative Complexity: Un Speeches on the Middle East Problem, 1947–1976. Journal of Conflict Resolution, 21(3), 427–442.

Suedfeld, P., Tetlock, P. E., & Streufert, S. (1992). 27 Conceptual/integrative complexity.

Tadmor, C. T., Galinsky, A. D., & Maddux, W. W. (2012). Getting the most out of living abroad: Biculturalism and integrative complexity as key drivers of creative and professional success. Journal of Personality and Social Psychology, 103(3), 520.

Tadmor, C. T., Tetlock, P. E., & Peng, K. (2009). Acculturation Strategies and Integrative Complexity The Cognitive Implications of Biculturalism. JOURNAL OF CROSS-CULTURAL PSYCHOLOGY, 40(1), 105-139.

Tadmor, C., & Tetlock, P. (2006). Biculturalism: A model of the effects of second-culture exposure on acculturation and integrative complexity. JOURNAL OF CROSS-CULTURAL PSYCHOLOGY, 37(2), 173-190.

Thies, C. G. (2009). The Conceptual Complexity of Central Bankers and the Asian Financial Crisis. POLITICAL PSYCHOLOGY, 30(3), 445-464.

Thoemmes, F. J., & Conway, L. G., III. (2007). Integrative complexity of 41 US presidents. POLITICAL PSYCHOLOGY, 28(2), 193-226.

Triandis, H. C. (2001). Individualism and Collectivism: Past, Present, and Future.

Van Swol, L. M., Carlson-Hill, C. L., & Lewis, E. E. A. (2018). Integrative Complexity, Participation, and Agreement in Group Discussions. SMALL GROUP RESEARCH, 49(4), 409-428.

Waterman, A. S. (1981). Individualism and interdependence. American Psychologist, 36(7), 762.

Wong, E. M., Ormiston, M. E., & Tetlock, P. E. (2011). The effects of top management team integrative complexity and decentralized decision making on corporate social performance. Academy of Management Journal, 54(6), 1207–1228.

Zhao, N., Jiao, D., Bai, S., & Zhu, T. (2016). Evaluating the Validity of Simplified Chinese Version of LIWC in Detecting Psychological Expressions in Short Texts on Social Network Services.

PLOS ONE, 11(6), e0157947.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2020). A comprehensive survey on transfer learning. Proceedings of the IEEE, 109(1), 43–76.

Zaib, M., Sheng, Q. Z., & Emma Zhang, W. (2020, February). A short survey of pre-trained language models for conversational ai-a new age in nlp. In Proceedings of the Australasian computer science week multiconference (pp. 1-4).